Hothersall, B., Whistance, L., Zedlacher, H., Algers, B., Andersson, E., Bracke, M., ... Main, D. (2016). Standardising the assessment of environmental enrichment and tail docking legal requirements for finishing pigs in Europe. *Animal Welfare*, *25*(4), 499-515. https://doi.org/10.7120/09627286.25.4.499

Standardising the assessment of environmental enrichment and tail docking legal requirements for finishing pigs in Europe

B Hothersall[1], L Whistance[1], H Zedlacher[2], B Algers[3], E Andersson[3], M Bracke[4], V Courboulay[5], P Ferrari[6], C Leeb[2], S Mullan[1], J Nowicki[7], M-C Meunier-Salaün[8], T Schwarz[7], L Stadig[9], D Main[1]

[1]School of Veterinary Science, University of Bristol, UK; [2]Department of Sustainable Agricultural Systems, University of Natural Resources and Life Sciences (BOKU), Austria; [3]Department of Animal Environment and Health, Swedish University of Agricultural Sciences (SLU), Sweden; [4]Wageningen Livestock Research, Wageningen University and Research Centre, The Netherlands; [5]IFIP Institut du Porc, France; [6]Centro Ricerche Produzioni Animali, Italy; [7]Department of Swine and Small Ruminants Breeding, University of Agriculture in Krakow, Poland; [8]Institut National de la Recherche Agronomique (INRA), France; [9]Animal Sciences Unit, Institute for Agricultural and Fisheries Research (ILVO), Belgium

Abstract

An online training package providing a concise synthesis of the scientific data underpinning EU legislation on enrichment and tail docking of pigs was produced in seven languages, with the aim of improving consistency of professional judgements regarding legislation compliance on farms. In total 158 participants who were official inspectors, certification scheme assessors and advisors from 16 EU countries completed an initial test and an online training package. Control group participants completed a second identical test before, and Training group participants after, viewing the training. In Section 1 of the test participants rated the importance of modifying environmental enrichment defined in nine scenarios from 1 (not important) to 10 (very important). Training significantly increased participants' overall perception of the need for change. Participants then rated nine risk factors for tail biting from 1 (no risk) to 10 (high risk). After training scores were better correlated with risk rankings already described by scientists. Scenarios relating to tail docking and management were then

described. Training significantly increased the proportion of respondents correctly identifying that a farm without tail lesions should stop tail docking. Finally, participants rated the importance of modifying enrichment in three further scenarios. Training increased ratings in all three.

The pattern of results indicated that participants' roles influenced scores but overall the training improved 1) recognition of enrichments that, by virtue of their type or use by pigs, may be insufficient to achieve legislation compliance, 2) knowledge on risk factors for tail biting and 3) recognition of when routine tail docking was occurring.

Keywords: animal welfare, enrichment, inspector, legislation, pig, tail docking,

Introduction

Animal welfare legislation has been developed for many countries and many species. However, the impact of legislation on animal welfare depends upon its full implementation in practice. In addition to appropriate awareness in the farming community, full implementation of EU legislation requires consistent assessment by those responsible for ensuring compliance. This can include official inspectors responsible for enforcement actions and assessors working for voluntary certification schemes that also aim to ensure compliance with legal prescriptions.

The complexity of the technical interpretation of legislation varies considerably between different requirements, depending on the availability of measurable criteria to define them. For example, assessing compliance with space allowance requirements necessitates measurement of the space, the number and, often, size of the animals housed in that space. In

comparison, environmental enrichment is more difficult to quantify and calls for a professional judgement. Standardising this professional judgement is necessary for consistent implementation. This can be particularly challenging when legislation, such as European Directives, is implemented by many different countries each using different inspection regimes.

This study describes an initiative aimed at improving the consistency of professional judgements needed to assess compliance with the environmental enrichment and tail docking requirements for finishing pigs included in EU Directive 2008/120/EC as detailed below:

"…pigs must have permanent access to a sufficient quantity of material to enable proper investigation and manipulation activities, such as straw, hay, wood, sawdust, mushroom compost, peat or a mixture of such, which does not compromise the health of the animals." *Annex 1, para 4*

"Neither tail-docking nor reduction of corner teeth must be carried out routinely but only where there is evidence that injuries to sows' teats or to other pigs' ears or tails have occurred. Before carrying out these procedures, other measures shall be taken to prevent tail-biting and other vices, taking into account environment and stocking densities. For this reason inadequate environmental conditions or management systems must be changed." *Annex 1, para 8*

These requirements are based upon extensive welfare research on tail biting in pigs (EFSA 2007a, EFSA 2014). The legislation is clearly intended to ensure pigs are provided with sufficient resources to satisfy their behavioural needs, to minimise the risks of injurious tail biting and to avoid unnecessary tail docking. The interaction between the various management factors is complex. The EFSA (2007a) scientific opinion concluded that: *"The*

occurrence of tail biting has a multi-factorial origin and there is evidence in the report that some causal factors have more weight, such as the absence of straw, the presence of slatted floors and a barren environment" and that "*there is little evidence that provision of toys such as chains, chewing sticks and balls can reduce the risk of tail biting."* The complex nature of the issue has led the UK Farm Animal Welfare Council (2009) to describe the presence of an intact uninjured tail on a growing pig at slaughter as an "*iceberg indicator*" because it may "*effectively summarise many measures of welfare and is easy to understand."* The report even suggested that an intact tail indicates that the *"animal's husbandry and management were of high quality and its welfare was good"*. In addition to the welfare implication, reducing the number and severity of tail lesions can have management benefits: tail biting and tail lesions have been associated with carcass condemnation (Valros et al 2004, Harley et al 2014).

Assessment of compliance with the enrichment and tail docking regulations requires consideration of both resource and animal-based outcomes. Determining the suitability of *"material to enable proper investigation and manipulation activities*" requires an assessment of both the substrate (resource) and the pig's behaviour (outcome). Where pigs have been tail docked an assessment of the *"evidence that injuries to .....tails have occurred"* (outcome) and of the changes to "*inadequate environmental conditions or management systems"* (resource) must also be made. The professional judgements associated with the assessment, therefore, necessitate considerable knowledge and understanding of the relevant scientific literature and its practical application.

The Food and Veterinary Office (FVO) is responsible for assessing compliance by each European Union Member State (MS). It observed that the enrichment and tail docking legislation has been inconsistently implemented in many states. Between January 2010 and August 2012, eight FVO mission reports included a specific recommendation concerning

inadequate implementation of these requirements (countries: IT, SK, HU, BE, PT, AT, DK, CZ) and 4 reported insufficient implementation (countries: RO, BG, IT (in 2010), FR). Only 2 missions reported compliance (countries: SE, LU) (Edman 2014). Campaign groups have also pressed for the implementation of environmental enrichment and tail docking requirements (CIWF 2015, Eurogroup for Animals 2012). In response to FVO concerns some Competent Authorities have developed guidance notes and training for Official Veterinarians (OVs) carrying out inspections (e.g. Ministerie van Landbouw Natuur en Voedselkwaliteit 2012). In some countries industry organisations have also been involved in interpreting the requirement (e.g. BPEX 2014). More recently, partly influenced by the current study, the European Commission is producing more detailed guidance on the interpretation of the directive 2008/120/EC.

Previous studies have shown that short educational interventions can improve knowledge of pig management and welfare and positively affect participant attitudes and behaviour (Hemsworth, et al. 1994, Coleman, et al. 2000,). Wright et al (2009) also found that formative online assessment of case studies improved vet students' ability to assess clinical signs of pig health and welfare. The present study describes the development and initial use of an online training package aimed at improving the consistency of the interpretation of the environmental enrichment and tail docking requirements included in EU Directive 2008/120/EC. The project was undertaken as part of a larger initiative, EU WelNet, to improve compliance with various aspects of animal welfare legislation. The remit of the training package was to summarise the welfare science basis for the legislative requirements for finishing pigs concerning environmental enrichment and tail docking. It included information on the motivation for tail biting and oral behaviour; risk factors for tail biting; attributes of effective enrichment material; tail docking practice, and welfare outcomes (oral behaviour and tail lesions).

The evaluation of the training package aimed to explore:

1) The extent to which participants took the type of enrichment and pig behaviour (manipulation of enrichment) into account when deciding whether material enables "proper investigation and manipulation activities".

2) Whether the on-line training changed understanding of the following aspects of the Directive: the relative risks of different management practices for tail biting; legal requirements needed to permit tail docking; the attributes of adequate enrichment to comply with legislation.

3) The effects of participants' professional roles (Official Inspector, Certification Assessor, Farm Advisor or Other) on their responses and whether training resulted in harmonisation of professional judgement (indicated by decreased variability in scores).

4) The views of the participants on the usefulness of the training package.

Materials and methods

Development of training package

The materials were produced collaboratively by a group of EU WelNet welfare scientists (the authors and their colleagues), who reviewed existing material including industry guidance, defined a preferred format and agreed the content (available at http://euwelnet.hwnn001.topshare.com). The training comprised a concise synthesis of scientific data underpinning EU legislation on enrichment and tail docking of finisher pigs. It took approximately 30 minutes to read and was designed to be attractive and accessible with illustrations, diagrams and video files to illustrate key points. The group also consulted with an advisory board, revised the content accordingly and defined an evaluation approach. The advisory board consisted of Chief Veterinary Officers of 27 EU Member States plus Croatia,

147 Norway and Switzerland, EU institutions involved with animal welfare (DGSANCO, EFSA,
148 FVO); International organisations (OIE, FAO, EUROFAWC); European organisations
149 representing animal and meat industries (Copa-Cogeca, EFFAB, IFAH, UECBV); Veterinary
150 and welfare science (FVE, ISAE, ISAH); Welfare education (ECVPH, ECAWBM); Welfare
151 organisations (CIWF, Eurogroup, FourPaws, World Animal Protection) and the European
152 Animal Welfare Platform (FAI). During development, a draft version of the tool was
153 distributed to the board and 15 responses were received from NGO, competent authorities,
154 science and industry groups from at least eight countries and three EU groups.

155 Feedback received from the advisory group contained many positive comments. The package
156 was generally considered a useful collation of the science, in an attractive, user-friendly
157 format and suitable for official inspectors. While it explored the relative merits of enrichment
158 materials, the package was not intended to attempt to define their absolute acceptability in
159 terms of compliance with EU legislation. This led to mixed responses from the advisory
160 board. A section on the limitations of different enrichment objects was positively received by
161 some, but others highlighted underlying uncertainties in interpreting the legislation, including
162 the need for clearer official guidance. ,. Some questioned whether the training should also
163 propose practical solutions for fully housed intensive production systems, or they indicated
164 perceived conflicts between science and practice regarding the utility of specific enrichments.
165 Within both the project team and the advisory board, it was considered difficult to find an
166 ideal format for each target audience, and to reconcile different opinions on key technical
167 issues such as the value of natural foraging behaviour or of enrichment objects and straw.
168 Another challenge was the different versions of the EU Directive amongst member states. For
169 example, the German and Polish translations use the word *movement* rather than
170 *manipulation* of materials, which have different connotations. These difficulties were
171 resolved as far as possible by an iterative process of revision and discussion, after which the

team translated the final package into English, Dutch, French, German, Italian, Polish and Spanish and recruited participants. The target audience consisted of official veterinarians responsible for assessing compliance with EU legislation within each country (AT, BE, DE, FR, IT, ES, PL, NL, UK,); certification scheme assessors and farm advisors (veterinary surgeons and other advisors). Pig producers did not form part of our final sample, though the tool was also considered suitable for future use with this group. The English language version was also made available to participants in Sweden, who were able to record their responses in Swedish.

Study design

All participants were invited to complete a demographic questionnaire and were then assigned to the Training or Control group. Assignation was automated and alternated between the two groups, including alternation within 'professional role'. All participants were then asked to complete a 27-question online test (Supplementary file S1, and described in the **Analysis** section) twice. Seven days after completing the first iteration of the test, participants were invited to log in again. Control group participants were then presented with the second iteration of the test (identical to the first) followed immediately by access to the training package, whereas Training group participants were directed to the training package and immediately after reaching the final page, to the test.

Feedback questionnaire

Following the second iteration of the test participants were asked to complete a feedback questionnaire consisting of nine statements (listed in Table 2) concerning the effect of the training on their confidence and understanding of EU legislation relating to tail docking and

enrichment. Participants could rate their agreement with each statement on a 10 point scale with anchors at 1 (no agreement) and 10 (full agreement).

Analysis

Quantitative analysis

Except where otherwise specified, random-intercept nested models were generated in MLwiN v2.25 for each question. The random effects were specified as Test iteration (1 or 2) as Level 1, nested within Participant (Level 2) within Country (Level 3). This multi-level structure allowed us to adjust for non-independence due to clustering within groups - for example, the tendency of an individual participant to give high scores, or of those from a particular country to score low. The influence of the following variables (fixed effects) on participants' scores were evaluated: test iteration (1 or 2), the professional role of the participant (Official Inspector, Certification Assessor, Farm Advisor or Other) and the interaction between group (training or control) and iteration. Inclusion of Iteration (1 or 2) as both a fixed and a random effect meant that it was treated as a repeated measure; the 'training x group' interaction was used to identify a divergence in scores between the groups following training, which would indicate a significant effect of training. The significance of individual predictors in a model was tested using Z-tests, whereby the coefficient was divided by the standard error of coefficient to generate respective Z-values. P-values were calculated as the area of the normal distribution greater than or equal to the Z-value, multiplied by two (two-tailed analysis). The significance of interactions in a model was tested using χ^2 tests and the deviance in log-likelihood between models both with and without the interaction. Data were transformed as necessary and standardised residuals were calculated and plotted to ensure that assumptions of normality and homoscedasticity were met.

In Section 1 of the test participants were presented with nine scenarios (see File S1), and for each, they were asked to rate the importance of modifying the enrichment in order to comply with EU legislation, from 1 (not important) to 10 (very important). Analysis examined the influence of various factors on the score given. Eight of the nine questions (scenarios) in this section were identical to another question except for the type of enrichment object present (clean, dry straw; wet, dirty straw; wood; or a chain), or the presence or absence of manipulation of the object (i.e. the scenario specified whether or not pigs were manipulating the object/s). For this section, a slightly different structure was applied: a single model was generated using data from the eight 'paired' questions and Question (Level 1) was nested within Participant (Level 2) nested within Country (Level 3). Enrichment object type and presence/absence of manipulation were added as fixed effects in addition to the fixed effects listed above.

Section 2 focussed on knowledge of tail biting risk factors. Nine risk factors were listed (File S1) and participants rated the level of risk from 1 (no risk) to 10 (high risk). We examined non-parametric correlations (Kendall's Tau) between the ranking of risk calculated by EFSA (2007a) and the ranking given by participants at each of the two time points. Tied ranks were assigned if values given by EFSA (2007a) were equal. One risk factor in our test ('pigs of different breeds within a group') was not mentioned in the EFSA chart; it was included for comparison purposes and was assigned the rank of zero (no risk). We then further examined the changes observed by creating nested models for each individual question.

Section 3 tested participants' knowledge of the legal requirements needed to permit tail docking. It described four scenarios relating to tail docking and management (File S1) and participants selected the action required from four options: none; identify and make suitable management changes; stop tail docking or permit tail docking. Participants could choose more than one answer. Responses were initially re-coded as a binary variable: correct (only

244 the correct answer selected) or incorrect (one or more wrong answers selected, including
245 where the participant *also* selected the correct answer) for each question. McNemar's tests
246 were conducted separately for the Control and Training groups to examine the change in the
247 proportion of correct answers between the two iterations.

248 In Section 4, participants were presented with three further scenarios (File S1), this time
249 focussing on the relationship between tail biting behaviour and the attributes of the
250 enrichment provided. They were again asked to rate the importance of modifying enrichment
251 in each case in order to comply with EU legislation. This was again analysed by nested
252 models of each question but respondents were also asked to give their own opinion on
253 appropriate action (see 'Qualitative analysis' section).

254 Finally, to examine whether training resulted in harmonisation of participants' views, we
255 calculated values for Levene's test of equality of variance for each of the variables where the
256 group x training interaction proved significant. The test was calculated once at each iteration
257 and compared Control and Training group participants' scores.

258 After completing the second iteration of the test, participants were invited to provide
259 feedback on the usefulness of the training package by indicating their agreement (0 = no
260 agreement; 10 = full agreement) with 9 statements.

261 ***Qualitative analysis***

262 Within Sections 4 and 5 free text questions provided participants with an opportunity to
263 express their own understanding of the legislation (File S1). After rating the importance of
264 modifying enrichment in each scenario in Section 4, participants were asked what they would
265 do next on this farm in their current role. It was then pointed out that different countries and
266 organisations interpret the EU directive differently; participants were further asked what they
267 would do if they were free from any such interpretations, and why. Section 5 presented two

268 further scenarios and free text questions. The first asked what the participant would do next in
269 this situation. They were then given additional details about environmental, management or
270 behavioural factors and asked what they would now do next.

271

272 **Results**

273 In total 158 participants (76 Control and 82 Training) from 16 countries completed both
274 iterations of the test, including 83 official inspectors, 23 certification scheme assessors and 24
275 farm advisors. The remaining 28 respondents were classified as 'Others'.

276 ***Quantitative analysis***

277 All results report predicted means ±SEM unless otherwise stated.

278 *Section 1: importance of modifying enrichment ('paired' pooled data from 8 scenarios)*

279 The type of enrichment object had significant effects on participants' scores, as did
280 presence/absence of evidence that pigs were manipulating the object/s. For object type
281 ($p=<0.001$) all four types differed from each other. Participants gave the lowest scores (least
282 important to modify the enrichment) for clean, dry straw followed by wood, followed by a
283 chain, and the highest scores (greatest need for change) where the question specified that wet
284 and dirty straw was present. Manipulation of those materials by pigs decreased scores by 0.80
285 (±0.12) ($p<0.001$). Scores were also influenced by participants' professional role ($p<0.001$),
286 with Official Inspectors and Others scoring significantly higher than Farm Advisors.
287 Certification advisors' scores were intermediate and did not differ significantly from any of
288 the other categories.

The interaction between Group and Iteration was also significant ($p<0.001$). This reflected very similar scores for the Control and Training groups at Iteration 1, with scores increasing (greater need for change) in the Training group only at Iteration 2 (**Figure 1**).

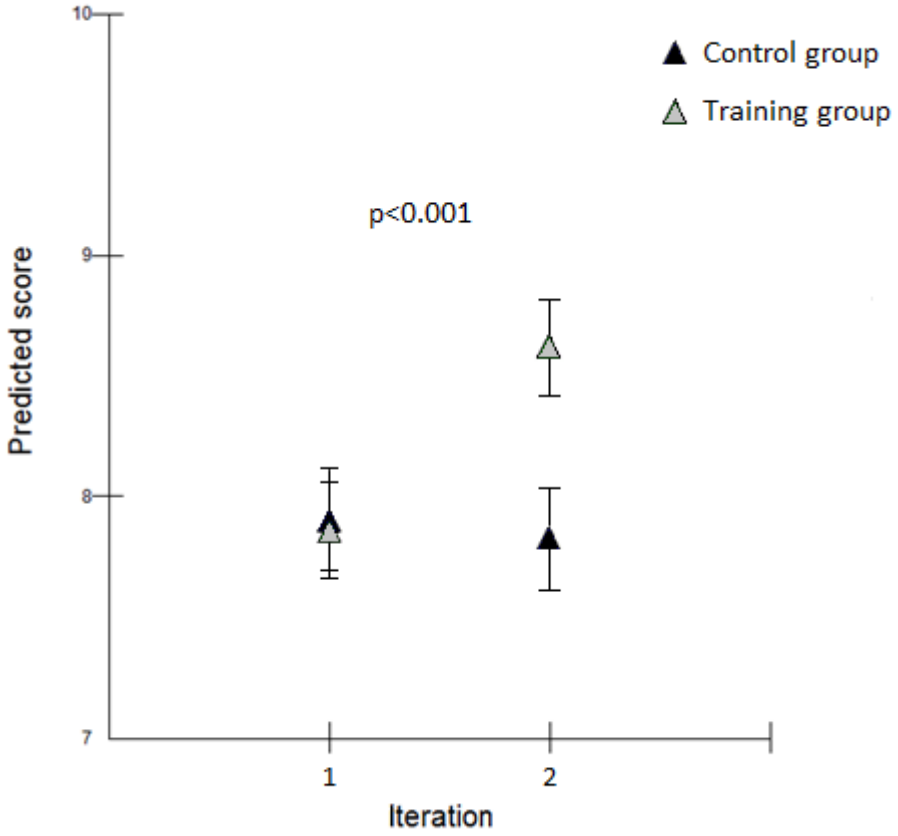


Figure 1: Mean group scores ± SEM for pooled questions in Section 1. Participants were given descriptions of enrichment and its use by pigs, and asked to score, on a scale from 0 to 10, the importance of modifying the enrichment provided, in order to comply with legislation. Scores are shown for Iteration 1 (before training) and Iteration 2 (after training). A higher score indicates a greater need for change.

Section 2: Risk factors for tail biting

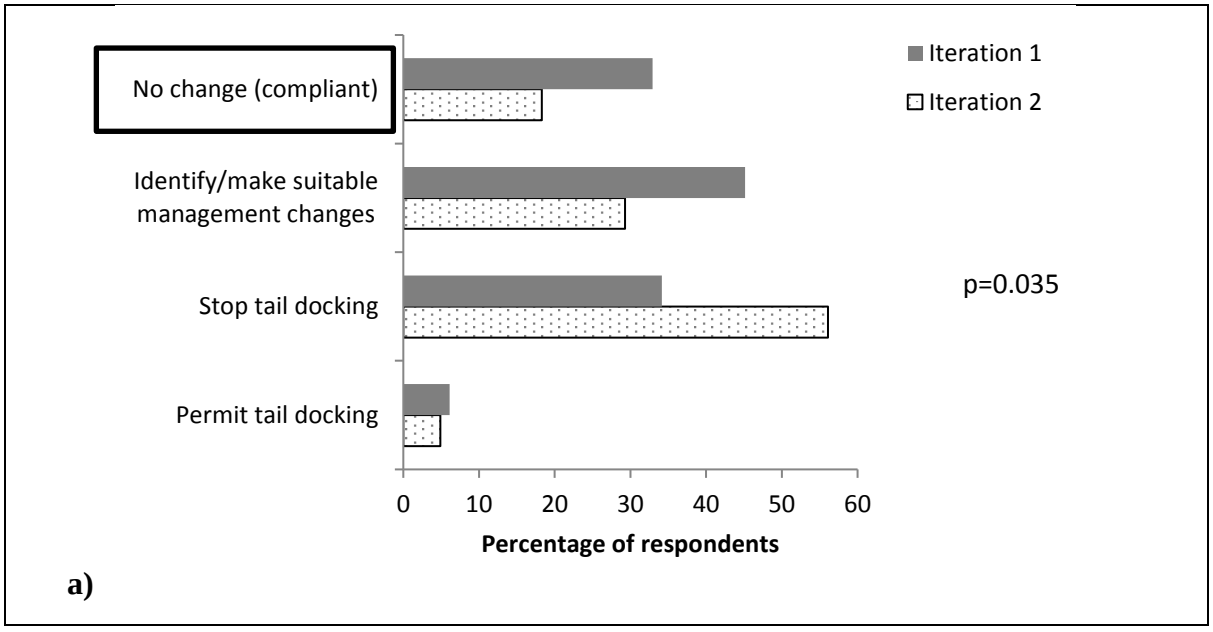
At Iteration 1, participants' ranking of risk factors for tail biting were already significantly correlated with the ranks shown in the EFSA (2007a) table (Training group and Control

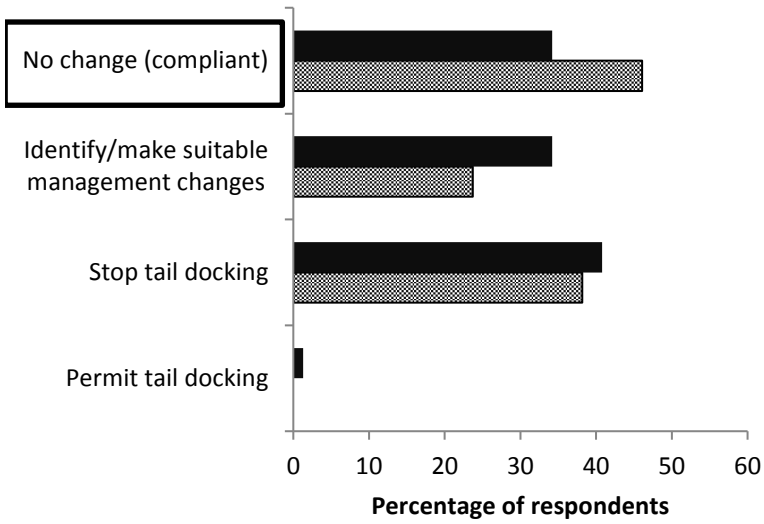
301 group, both Kendall's Tau = 0.572; p=0.035). At Iteration 2 the correlation explained more of
302 the variability and was more significant in both groups, but the change was greater in the
303 Training group (Control group, Tau = 0.629, p=0.02; Training group, Tau = 0.800, p=0.003).
304 Modelling of the data for individual questions revealed that this was due to a significant
305 Group x Iteration interaction for three of the variables. Participants rated the risk posed by a
306 'barren environment' as very high in the initial test but this nonetheless increased
307 significantly after training (Iteration 1: Training group 8.96; Control group: 9.28. Iteration 2:
308 Training group 9.54; Control group 9.08; p=0.002). Conversely, training led to moderate
309 *decreases* in risk ratings for 'heat stress' (from 7.90 to 6.84 in the Training group, whereas
310 Control group scores went from 8.04 to 8.22; p=0.0003) and 'high stocking density' (from
311 8.95 to 8.18 in the Training group; from 9.07 to 8.91 in the Control group; p=0.005). The
312 latter decreases were in accordance with the information provided during training.
313 Professional role did not significantly affect scores for any of these three variables. The
314 variable 'pigs of different breeds within a group' (which is not a recognised risk) was not
315 mentioned in the training and training did not significantly affect scores for this variable
316 either.

317 *Section 3: Tail biting and management practices - identification of action required to achieve*
318 *compliance (four scenarios)*

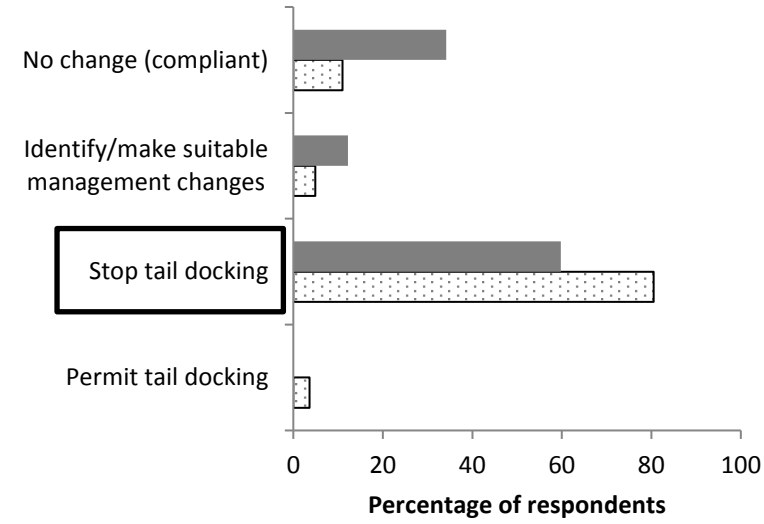
319 There was no significant change between Iterations 1 and 2 in the proportion of Control
320 Group participants identifying the correct action in any of the four scenarios about tail biting
321 and management practices. In contrast, Training group responses changed significantly in
322 two of the scenarios. In Question 19, following training there was a significant *decrease*
323 (McNemar's test; p=0.035), from 33% to 18% in the percentage of participants correctly
324 recognising that no action was required at a farm which had made suitable management
325 changes after a recent tail biting outbreak (**Figure 2a**; control group responses are shown for

comparison in **Figure 2b**). The scenario stated that there were no pigs with fresh tail lesions but that pigs with healed tail lesions were present. Following training, 56% of Training group participants incorrectly answered that the farm should stop tail docking, compared with 34% in Iteration 1. In Question 20, the proportion of respondents correctly identifying that a farm with no evidence of tail lesions should stop tail docking increased significantly from 60% to 80% after training (McNemar's test; p=0.001) (**Figure 2c**; control group responses **Figure 2d**). No significant improvement was seen in either of the remaining questions, where correct scores were already very high at Iteration 1: over 80% in Question 21 and over 90% in Question 22 (see supplementary files S1 (questionnaire) and S2 (results summary).

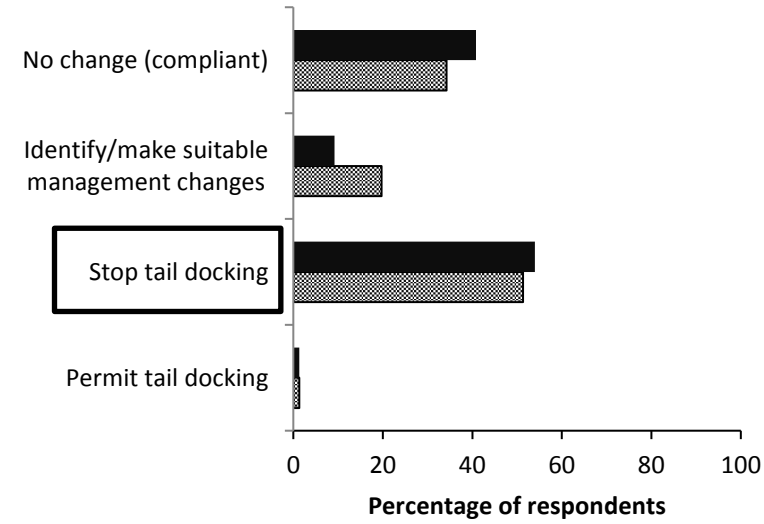



No change (compliant)
Identify/make suitable management changes
Stop tail docking
Permit tail docking
Iteration 1
Iteration 2
0
10
20
30
40
50
60
Percentage of respondents

b)


No change (compliant)
Identify/make suitable management changes
Stop tail docking
Permit tail docking
Iteration 1
Iteration 2
p=0.001
0
20
40
60
80
100
Percentage of respondents

c)


No change (compliant)
Identify/make suitable management changes
Stop tail docking
Permit tail docking
Iteration 1
Iteration 2
0
20
40
60
80
100
Percentage of respondents

d)

Figure 2: Distribution of answers in Section 3. Scenarios relating to tail biting and management practices were described. Respondents were asked to identify the action required to achieve compliance. They could select more than one answer. The correct answer is marked by a black outline. Question 19: All pigs are tail docked. A recent outbreak of tail biting has occurred; suitable management changes have been made in response. Pigs with healed tail lesions are present. No pigs with fresh tail lesions are present. **a)** training group; **b)** control group. Question 20: All pigs are tail docked. Current management practices are suitable. No pigs with healed lesions and no pigs with fresh tail lesions are present. **c)** training group; **d)** control group.

Section 4: Importance of modifying enrichment (three scenarios)

In all three of the scenarios in Section 4 there were significant Training Group x Iteration interactions, indicating that training increased scores by between 0.29 and 1.3; there were also significant effects of the professional role of the participant (**Table 1**).

The first scenario (Question 23) described a barren environment. Official inspectors scored higher (greater need to modify enrichment) than Farm Advisors; other pairs of roles did not differ significantly in their scores. At Iteration 2, scores increased slightly in the training group and decreased slightly in the control group. In the second Scenario (Question 24), tail lesions were present and pigs were provided with but not manipulating straw that was wet and dirty. Official Inspectors again scored highest, giving significantly higher ratings than either Farm Advisors or Other. The significant Training Group x Iteration interaction represented an increase in scores at Iteration 2 in the training group only.

For both of these questions, the majority of respondents gave the maximum score of 10. This resulted in means for all Group/Iteration combinations of ≥ 9.52 and 8.29 for Questions 23 and 24 respectively. As a result, the data could not be satisfactorily transformed and model estimates may be imprecise, so should be interpreted with caution.

In the final scenario (Question 25) pigs were able to reach two chains but were only manipulating one; there were no pigs with tail lesions. All groups scored a lesser need for change (lower score) in this scenario, and in contrast with the previous scenarios, scores of Official Inspectors (5.20±0.58) and Farm Advisors (4.12±0.72) overlapped and were both significantly *lower* than those of Certification Assessors (6.63±0.73) or Others (6.53±0.72), which did not differ from one another. Again, the significant Training Group x Iteration interaction represented an increase in scores at Iteration 2 in the training group only.

Table 1: Model predictions for the effects of Training Group x Iteration interaction and professional role in Section 4. Participants were presented with farm scenarios and asked to score the importance of modifying the enrichment provided, in order to comply with legislation. A higher score indicates a greater need for change. Co-efficients represent the predicted change in score relative to the reference categories specified in the model (here: Iteration = 1, Training group = Control; Role = Official Inspector). Roles differing from Official Inspector are shown in bold; a negative co-efficient indicates scores were lower, and a positive one higher.

Question	Training x Iteration interaction		Professional role[1]	
	Co-efficient (SE)	p-value	Co-efficient (SE)	p-value
Q23: Barren environment; no substrate or	+0.29 (0.10)	0.006	CA: -0.27 (0.15) Other: -0.23 (0.14)	0.02

enrichment; tail lesions present			**FA: -0.40 (0.14)**	
Q24: Wet, dirty straw; no manipulation of straw; tail lesions present	+0.57 (0.22)	0.01	CA: 0.28 (0.32) **Other: -0.65 (0.30)** **FA: -0.74 (0.31)**	0.05
Q25: Two chains in reach; one manipulated; no tail lesions present	+1.31 (0.44)	0.003	**CA: 1.43 (0.59)** **Other: 1.33 (0.55)** FA: -1.09 (0.57)	0.0006

[1] CA = Certification Assessor; FA = Farm Advisor. SE = Standard Error

Equality of variance as a measure of harmonisation

For each of the questions where the group x training interaction proved significant, Levene's test of equality of variance was calculated. This compared Control and Training group participants' scores and was calculated once at each iteration.

At Iteration 1, variance was significantly different for only 1 of 11 questions: Section 4, Scenario 1 (importance of modifying enrichment: barren environment; tail lesions present. f=4.65, p=0.033), where variance was greater within the Training group than within the Control group.

At Iteration 2, variance differed significantly between Training and Control groups for 10 of the 11 questions. In 8 of these 10 cases, variance was now lower in the Training group than the Control group, indicating that training had reduced variability in the scores given (i.e. harmonised judgement). The 2 cases in which variance was higher in the Training group were Q2 (heat stress) and Q9 (high stocking density) in the 'risk factors' (Section 2). The multi-level analysis had previously revealed that training group scores for these two risk factors decreased after training, consistent with the information provided in training.

396 ***Qualitative analysis***

397 In Section 4 (3 scenarios exploring the importance of modifying enrichment) and Section 5
398 (additional 2 detailed scenarios followed by further information) participants were asked
399 'what would you do?' The aim of the questions was to encourage participants to describe
400 their own professional response to the scenarios. Unfortunately most respondents reported
401 the actions that should be taken by the farmer (e.g. 'check ventilation' or 'provide more
402 enrichment') making it difficult to analyse the impact of the training tool on the participants.
403 Even though respondents were not specifically asked to provide their opinion on the
404 enforcement of legislation, some comments were included in the question about their role.
405 There was a full range of suggestions such as arguing for allowing or banning routine tail
406 docking, permitting more flexible interpretation to suit the local conditions or suggesting
407 greater harmonisation of enforcement.

408 ***Feedback questionnaire***

409 In total 150 of the participants completed the feedback questionnaire for the training tool
410 (**Table 2**). Responses indicated that it was very well received, with mean scores of at least
411 7.49/10 for all questions. The highest mean score (8.81) was given in response to the
412 statement asking whether participants would recommend the tool to others involved in
413 assessing finisher pig welfare. The lowest mean score (7.49) was given for the better
414 understanding of production losses associated with tail biting.

415

416 **Table 2:** Summary of feedback on training tool from 150 participants. Score: 1 = no
417 agreement to 10 = full agreement**.** Statements are ordered by decreasing mean.

Question No.	Please indicate your level of agreement with the following statements.	Mean (SD)

9	I would recommend this training tool to other persons involved in assessing the welfare of finisher pigs in the EU.	8.81	(1.89)
3	The training has increased my understanding of which enrichment materials best enable the expression of 'proper investigation and manipulation activities'.	8.16	(2.25)
6	The training has increased my understanding of the EU legislation related to enrichment provision.	7.85	(2.33)
4	The training has increased my understanding of the relationship between tail biting and housing and management practices.	7.83	(2.14)
2	The training has increased my understanding of the relationship between tail biting and a lack of opportunity to express foraging behaviour and other investigation and manipulation activities.	7.83	(2.29)
1	The training has increased my understanding of the importance for pigs to be able to express foraging behaviour and other investigation and manipulation activities.	7.82	(2.39)
8	The training has increased my confidence in interpreting EU legislation on farms.	7.71	(2.43)
7	The training has increased my understanding of the EU legislation related to tail docking of pigs.	7.48	(2.67)
5	The training has increased my understanding of the relationship between tail biting and production losses.	7.46	(2.47)

Discussion

The development of the training package demonstrated that a large group of welfare scientists were able to work collaboratively and incorporate diverse viewpoints to create a product that was received well by both the advisory board (as described in the 'Development of tool' section) and the intended audience (according to their feedback). The evaluation provided encouraging confirmation that participants had a good general understanding of the legislation and took the attributes and use (manipulation) of enrichment into account when

assessing compliance. Moreover, the training package had a significant positive influence on participants' understanding of the relevant legislative requirements, the importance of modifying certain enrichments and of certain tail biting risk factors. While training did not influence every individual question, results consistently indicated that it improved participants' understanding of situations where compliance was relatively difficult to assess. Professional judgements were also consistently harmonised when assessed immediately after training, measured by a reduction in variability of scores. No such pattern of changes was seen for Control group participants. Responses to the feedback questionnaire were very positive, with respondents particularly reporting that they would recommend the training package to colleagues. Indeed, following initial dissemination of the results, the project team has received interest and requests to use the training package from a number of sources.

In Section 1, participants took enrichment properties and use into account when assessing their suitability: they gave lower scores (less need to modify enrichment) when the objects or substrate were manipulated, and when clean, dry straw was given compared with other enrichments. This was in line with the scientific evidence (Moinard, et al. 2003, EFSA 2007b, Scott, et al. 2007, Studnitz, et al. 2007, Van De Weerd, et al. 2006) summarised in the training package. Training group overall scores (for the eight 'paired' questions) increased at Iteration 2, demonstrating that training was effective in increasing awareness of enrichments that would be less likely to comply with legislation. This suggests that a training package might help minimise variation in professional judgements of legislation compliance. However it is also probably reasonable to suggest that further official guidance on the principles of suitable enrichment and acceptability of specific common enrichments might have an even greater influence on reducing the variability in the assessment of compliance.

Encouragingly, participants' rankings of risk factors in Section 2 already correlated with those published by EFSA (2007a) prior to training. The strength and significance of the

452 correlation increased for both groups, suggesting that taking part in the test prompted
453 reflection. Nevertheless, the increase was greater for the Training group and analysis of
454 individual risk factors indicated that the changes were in accordance with the information
455 provided in training.

456 When participants were given scenarios and asked to identify the action needed to ensure
457 compliance (Section 3), significant changes were seen across the iterations for the training
458 group only. Knowledge was improved after training in one question, but the percentage
459 replying correctly actually decreased in another. The latter concerned a situation where tail
460 biting appeared to have stopped. The test describes two very similar but not identical
461 situations. It was intended to force participants to find those small differences in scenarios
462 that could be reflected in practice, but it failed. The problem could be of several origins
463 including construction of the training tool, general principles and imperfections in e-learning
464 systems, and last but not least the individual perception of participants. Throughout the
465 training tool, a large amount of information was given about the legislative, biological and
466 environmental background of tail biting and tail docking. Some situations describing
467 compliance or the lack of compliance with the legislation were not stated until the last section
468 of the training tool, leading perhaps to decreased attention of participants and thus an
469 incorrect interpretation of some scenarios described in the test. Training may therefore have
470 engendered greater confidence that management changes would resolve the problem,
471 allowing the producer to cease tail docking straight away. This problem could also indicate a
472 wider imperfection of e-learning formats. Data from literature on distance learning show that
473 some technical aspects of an e-tool lead to uniformity, which can be an obstacle in
474 interpretations of dynamic situations, especially in heterogeneous environments (Birnbaum
475 2001). More broadly, Greatrix (2001) argues that standardising assessment cannot itself
476 guarantee that assessors are comparable and that those being assessed meet or even

understand the standards required of them. Some of the answers during the second iteration suggested that some participants applied the legislation simply and in a uniform way for decision making, without analysing individual farm situations. The increase in wrong answers to question 19 was very similar to the increase in correct answers to question 20, suggesting that some participants in both scenarios made a decision using a uniform scheme. No change was seen for the other two scenarios, where compliance or non-compliance was arguably more obvious. Most respondents answered these correctly at Iteration 1, leaving little room for improvement.

The pattern of improvement was again seen in Section 4, where scores in all three scenarios increased at the second iteration in the Training group but not in the Control group. This was despite already very high initial scores for both groups in the first (barren environment) and second (wet, dirty straw) scenarios. The third scenario described the provision of chains that were partly being manipulated and was probably the least obvious to assess as compliant or non-compliant.

Equality of variance tests provided evidence that training also harmonised professional judgements. Training and Control group participants were very similar in the variability of their answers at the first iteration. Of all the questions that were influenced by training, variance differed in only one out of 11 questions: Training participants were more variable. In contrast, at Iteration 2, variance differed in ten of the 11 cases. In eight cases, Training participants were now *less* variable than Controls, suggesting that – as intended – the effect of training was to make their assessments more similar. Interestingly, the opposite result was seen for the remaining two questions: for ‘heat stress’ and ‘high stocking density’ in the Section 2 risk factors, the groups did not differ significantly at Iteration 1 but the Training group was more variable after training. Modelling of responses to these individual questions indicated that mean ranks assigned to them decreased after reading the training package,

502 which presented participants with evidence that these were relatively low risk factors. The
503 increase in variance suggests that not all participants picked up on this, so future training
504 could elaborate on or emphasise such information further. This may also reflect the training
505 package's focus on environmental enrichment and future training tools could target additional
506 risk factors if desired.

507 Anneberg et al (2012) reported that Danish livestock producers perceived welfare inspectors
508 as 'outsiders' who as such are unable to make fair judgements of farms. The authors raised
509 the question of whether the authoritative position of inspectors discourages dialogue or
510 farmers' motivation to make improvements. Our survey included providers of advice and
511 guidance as well as those with a statutory remit, and our modelling structure allowed us to
512 investigate the effect of professional role on participants' responses. Role influenced scores
513 in Sections 1 and 4, though not in Section 2 (explored for individual risk factors whose scores
514 were influenced by training). In Section 1 and in two of the three scenarios of Section 4,
515 Official Inspectors gave higher scores (greater need to improve enrichment) compared with
516 Farm Advisors. However, in the third scenario (chains provided with 1 of 2 being
517 manipulated; no lesions) Official Inspectors gave scores as low, or lower, than the other
518 groups. It is plausible that Official Inspectors' role gave them greater authority or confidence
519 to require changes of producers, but the latter result suggests that they (more than other
520 groups) interpreted chains as compliant. If Official Inspectors commonly observe chains in
521 use as enrichment, it may be that they judge them as compliant *if* they appear to be effective
522 in preventing tail biting. Farm Advisors are likely to have similar experiences and their scores
523 for this scenario were also low; it is possible that Certification Assessors had reference to
524 additional scheme-specific criteria which encouraged them to view chains as non-compliant.

525 Unfortunately the methodology used in this paper did not produce a reliable description of the
526 participant's professional response whilst attending pig farms. It is suggested that in-depth

527 interviews and observing farm visits as used by Roe et al (2011) would be required to
528 understand the complex interaction between the legislative requirements, the farmer and the
529 role of assessor or advisor.
530 Since this EU legislation is controversial it is not surprising,, that respondents used the free
531 text option to comment on enforcement of legislation. Options to improve enforcement have
532 been discussed by Lerner and Algers (2013) and re-iterated in a recent report to the European
533 Parliament Committee on Petitions examining implementation of Directive 2008/120/EC
534 (Marzocchi 2014). A survey of European pig farmers reported a consensus that legislation
535 and regulation needed to be harmonised across nations to ensure a "level playing field",
536 tempered by concerns from some that selected welfare measures "may conflict with farmers'
537 definitions of animal welfare and good farming practices" (Bock & van Huik 2007). In
538 Anneberg et al's 2012 study, farmers stated a desire for a set of rules and mentioned the
539 importance of these in ensuring quality assurance. Yet the same individuals believed that
540 assessment of compliance should reflect the producer's individual situation, citing aspects
541 such as facilities, staffing, experience and the overall quality or productivity of the farm.
542 They also argued that inspections are subjective or inconsistent depending on an inspector's
543 personality, attitude or personal interpretation of the legislation (Anneberg, et al. 2012). If
544 true, this is worrying, but empirical data are scarce. Mullan et al (2011) found that attitude to
545 farm animal welfare did *not* confound training in pig welfare outcome measures in a group of
546 UK farm assurance assessors, but as borne out by our results, many factors may still
547 influence the response of an assessor once a problem is identified.
548
549 The changes brought about by training were modest, resulting in group differences of no
550 more than 2 in mean scores at Iteration 2. Training was intentionally brief, and a more in-
551 depth intervention might be needed to influence scores more strongly. Our focus was on

552 presenting a digest of scientific evidence, but the materials could easily be adapted to include
553 additional clarification of legislative requirements and even official guidance as part of a
554 wider initiative aimed at promoting compliance with the EU directive. For a number of
555 questions, the capacity for training to increase scores was limited because initial scores or
556 correct responses were already high, indicating that participants already perceived a great
557 need for change or had already identified the appropriate action. Inclusion of scenarios of
558 varying 'difficulty' was helpful to identify where participants needed additional guidance or
559 clarification of the legislation, but the ceiling effect may have been exacerbated by the use of
560 Likert items and a numeric response format, where it was possible to give many items high or
561 low ratings. The use of this format was considered carefully during development; it was
562 considered to reduce ambiguity or bias due to difficulties in translating option answers or
563 descriptors (e.g. Harris-Kojetin, et al. 1999). Use of agreement scales also maintained the
564 focus on users' perceptions of the legislation rather than attempting to provide definitive
565 interpretation of compliance or non-compliance in more debatable cases. Overall, the pattern
566 of results showed training helped participants to identify (or increased their perception of the
567 need to modify) enrichments that were less likely to achieve compliance. It was not possible
568 to monitor contact between participants following training or testing and we acknowledge
569 that discussion of the contents of the training package (or even simply dialogue about the
570 legislation) could have affected participants' score. It would be very difficult to avoid this
571 problem in any online, multi-country evaluation, but it is a potential confound that cannot be
572 quantified. Indeed, simply taking part in the study may have increased awareness of the
573 relevant legislation in both Control and Training group participants.

574 For logistical reasons, the evaluation was restricted to a short-term assessment immediately
575 after training. Foshay and Tinkey (2007) note that recall of knowledge attained through
576 training is likely to diminish over time, and that written tests examine declarative knowledge

577 without necessarily reflecting professional competence. In previous evaluations of a short
578 educational intervention (1 hour information session plus written/visual handouts
579 (Hemsworth, et al. 1994) and a supplementary small group session (Coleman, et al. 2000)
580 improvements in knowledge about pig husbandry and welfare had positive effects on the
581 attitude and behaviour of farm staff. It was beyond the scope of the current study to assess
582 whether training influenced the decisions or actions taken by assessors when subsequently
583 assessing farms. It was also recognised during the development of the project that the training
584 package would have limited impact directly on farmers, although was considered suitable for
585 them. Farmer-focused initiatives will also be needed to promote compliance with legislation.

586

587 **Conclusions**

588 The training package presented an attractive, accessible summary of the scientific basis for
589 the legislation on environmental enrichment and tail docking contained within EU Directive
590 2008/120/EC. The package was designed for professionals involved in the assessment of
591 finisher pig welfare and was well received by participants. Short-term evaluation indicated
592 that completing the training improved the consistency of participants' professional
593 judgements and improved knowledge of several aspects of the legislation, particularly where
594 assessment of compliance might be considered contentious or difficult. Participants strongly
595 agreed that they would recommend the training package to other persons involved in
596 assessing the welfare of finisher pigs in the EU.

597

598 **Animal Welfare Implications**

599 Animal welfare legislation is designed to afford some protection to animals and ensuring
600 appropriate enforcement is part of the way that this is achieved. Improving the understanding

601 of welfare legislation of official inspectors, inspectors of voluntary certification schemes and
602 farm advisors is the first step towards the goal of farmers making changes to achieve
603 legislation compliance and improve welfare. The training package proved a valuable tool in
604 this first step, however, further research is required to evaluate how inspectors' improved
605 understanding of the legislation affects their actions on farms and on subsequent changes
606 undertaken by farmers.

607

608 **References**

609 **Anneberg I, Vaarst M, and Sorensen JT** 2012 The experience of animal welfare
610 inspections as perceived by Danish livestock farmers: A qualitative research approach.
611 *Livestock Science 147:* 49-58.

612 **Birnbaum R** 2001 *Management fads in higher education. Where they come from, what they*
613 *do, why they fail.* Jossey-Bass: San Francisco

614 **Bock BB, and van Huik MM** 2007 Pig Farmers and Animal Welfare: A study of beliefs,
615 attitudes and behaviour of pig producers across Europe., In: Kjaernes U, Miele M and Roex J
616 (eds) *Attitudes of Consumers, Retailers and Producers to Farm Animal Welfare, Welfare*
617 *Quality Report no. 2,* pp 73-124: Lelystad, NL

618 **BPEX** 2014 Environmental Enrichment, Examples, British Pig Executive: UK.
619 http://www.bpex.org.uk/health-welfare/welfare/environment-enrichment/environmental-
620 enrichment-examples/. Accessed 12th March 2015

621 **CIWF** 2015 Project Pig, Compassion in World Farming: UK. http://www.ciwf.org.uk/our-
622 campaigns/pigs/. Accessed 12th March 2015

623 **Coleman GJ, Hemsworth PH, Hay M, and Cox M** 2000 Modifying stockperson attitudes
624 and behaviour towards pigs at a large commercial farm. *Applied Animal Behaviour Science*
625 *66:* 11-20.

626 **Edman F** 2014 Do the Member States of the European Union comply with the legal
627 requirements for pigs regarding manipulable material and tail docking? Swedish University
628 of Agricultural Sciences: Uppsala, Sweden

629 **EFSA** 2007a The risks associated with tail biting in pigs and possible means to reduce the
630 need for tail docking considering the different housing and husbandry systems. *The EFSA*
631 *Journal 611:* 1-13.

632 **EFSA** 2007b Animal health and welfare in fattening pigs in relation to housing and
633 husbandry. *The EFSA Journal 564:* 1-14.

634 **EFSA** 2014 Scientific Opinion concerning a multifactorial approach on the use of animal and
635 non-animal-based measures to assess the welfare of pigs. *EFSA Journal 12(5):* 3702

636 **Eurogroup for Animals** 2012 Danish member makes formal complaint to European
637 Commission about tail docking. Eurogroup for Animals: Brussels, Belgium

638 **FAWC** 2009 Farm Animal Welfare: Past, Present and Future. *Farm Animal Welfare*
639 *Committee:* London, UK

640 **Foshay WR, and Tinkey PT** 2007 Evaluating the effectiveness of training strategies:
641 Performance goals and testing. *Ilar Journal 48:* 156-162.

642 **Greatrix P** 2001 Quality assurance into the 21st century: Command and control or
643 enlightened accountability? *Perspectives 5:* 12-16.

Harley S, Boyle LA, O'Connell NE, More SJ, Teixeira DL, and Hanlon A 2014 Docking the value of pigmeat? Prevalence and financial implications of welfare lesions in Irish slaughter pigs. *Animal Welfare 23:* 275-285.

Harris-Kojetin LD, Fowler FJ, Brown JA, Schnaier JA, and Sweeny SF 1999 The use of cognitive testing to develop and evaluate CAHPS (TM) 1.0 core survey items. *Medical Care 37:* MS10-MS21.

Hemsworth PH, Coleman GJ, and Barnett JL 1994 Improving the attitude and behaviour of stockpersons towards pigs and the consequences on the behaviour and reproductive performance of commercial pigs. *Applied Animal Behaviour Science 39:* 349-362.

Lerner H, and Algers B 2013 Tail docking in the EU: a case of routine violation of an EU Directive., In: Röcklinsberg H and Sandin P (eds) *The Ethics of Consumption: the citizen, the market, and the law.* pp 374-378. Wageningen Academic Publishers: The Netherlands

Marzocchi O 2014 Routine tail-docking of pigs. *European Parliament Directorate General for Internal Policies:* Brussels, Belgium

Ministerie van Landbouw Natuur en Voedselkwaliteit 2012 Goede afl eiding voor uw varkens. *Ministerie van Landbouw Natuur en Voedselkwaliteit:* Netherlands

Moinard C, Mendl M, Nicol CJ, and Green LE 2003 A case control study of on-farm risk factors for tail biting in pigs. *Applied Animal Behaviour Science 81:* 333-355.

Mullan S, Edwards SA, Butterworth A, Whay HR, and Main DCJ 2011 A pilot investigation of Farm Assurance assessors' attitude to farm animal welfare as a confounding factor to training in pig welfare outcome measures. *Animal Welfare 20:* 413-421.

Roe E, Buller H, and J Bull 2011 The performance of farm animal assessment. *Animal Welfare 20*: 69-78.

Scott K, Chennells DJ, Armstrong D, Taylor L, Gill BP, and Edwards SA 2007 The welfare of finishing pigs under different housing and feeding systems: liquid versus dry feeding in fully-slatted and straw-based housing. *Animal Welfare 16:* 53-62.

Scott K, Taylor L, Gill BP, and Edwards SA 2007 Influence of different types of environmental enrichment on the behaviour of finishing pigs in two different housing systems - 2. Ratio of pigs to enrichment. *Applied Animal Behaviour Science 105:* 51-58.

Studnitz M, Jensen MB, and Pedersen LJ 2007 Why do pigs root and in what will they root? A review on the exploratory behaviour of pigs in relation to environmental enrichment. *Applied Animal Behaviour Science 107:* 183-197.

Valros A, Ahlstrom S, Rintala H, Hakkinen T, and Saloniemi H 2004 The prevalence of tail damage in slaughter pigs in Finland and associations to carcass condemnations. *Acta Agriculturae Scandinavica Section a-Animal Science 54:* 213-219.

Van De Weerd HA, Docking CM, Day JEL, Breuer K, and Edwards SA 2006 Effects of species-relevant environmental enrichment on the behaviour and productivity of finishing pigs. *Applied Animal Behaviour Science 99:* 230-247.

Wright AJ, Powney SL, Nevel A, and Wathes CM 2009 Pig Welfare Assessment: Development of a Protocol and Its Use by Veterinary Undergraduates. *Journal of Veterinary Medical Education 36:* 50-61.

Supplementary material

Supplementary file S1 (attached) – Test text.

Supplementary file S2 (below) – raw results.

Supplementary file S3 (attached) Abbreviations

Question	*Number in test*	*Description*		*Control group*		*Training group*	
				Mean	*S.D.*	*Mean*	*S.D.*
Section 1: In order to ensure compliance with the EU Directive 2008/120/EC, how important is it to modify the enrichment provided on these farms? For each scenario, choose a point on the scale ranging from 'not important to very important'.							
		Object	Manipulating object/straw?	Iteration 2 only			
S1Q1	*1*	None	Yes (dung only)	9.49	0.87	9.49	1.43
S1Q2	*2*	Wood	No	7.53	2.43	8.48	2.05
S1Q3	*3*	Wet / dirty straw	Some	8.17	1.81	8.83	1.59
S1Q4	*4*	Clean / dry straw	Some	3.82	2.64	3.83	2.68
S1Q5	*5*	Chain	Some	6.30	2.68	8.15	1.79
S1Q6	*6*	Clean / dry straw	No	4.42	2.88	4.70	3.06
S1Q7	*7*	Chain	No	7.47	2.64	8.77	1.89
S1Q8	*8*	Wood	Some	6.01	2.74	7.26	2.27
S1Q9	*9*	Wet / dirty straw	No	8.58	1.82	8.87	1.71

Section 2: Please rate the level of risk for each factor by choosing a point on the scale ranging from ‘no risk to high risk’.

		Description	Iteration 2 only			
S2Q1	*10*	Presence of tail-bitten animals	9.50	0.97	9.23	1.64
S2Q2	*11*	Heat stress	8.22	1.80	6.84	2.41
S2Q3	*12*	Mixing of animals (not including mixing at weaning)	8.45	1.48	7.65	2.05
S2Q4	*13*	Absence of natural light	6.16	2.33	5.85	2.64
S2Q5	*14*	Poor herd health	8.05	1.80	7.50	2.17
S2Q6	*15*	Sudden changes in diet	7.96	1.93	7.44	2.08
S2Q7	*16*	A barren environment (no substrate and no enrichment)	9.08	1.38	9.54	0.79
S2Q8	*17*	Pigs of different breeds within a group	6.07	3.03	5.45	2.79
S2Q9	*18*	High stocking density	8.91	1.30	8.18	1.87

Section 4: In order to ensure compliance with the EU Directive 2008/120/EC, how important is it to modify the enrichment provided on these farms? For each scenario, choose a point on the scale ranging from 'not important to very important'						
		Description	Iteration 2 only			
S4Q1	*23*	Barren , Tail lesions	9.70	0.83	9.83	0.54
S4Q2	*24*	Wet / dirty straw, No manipulation, Tail lesions	9.01	1.79	9.52	0.88
S4Q3	*25*	Chains, Some manipulation, No tail lesions	5.24	3.17	7.21	2.54

Section 3: Tick the appropriate option/s to indicate what action is needed to achieve compliance (% of respondents)									
Scenario			Iteration 1				Iteration 2			
			Control		Training		Control		Training	
			Wrong	Right	Wrong	Right	Wrong	Right	Wrong	Right
S3Q1	*19*	Compliant (tail-biting outbreak addressed adequately)	67.1	32.9	69.5	30.5	56.6	43.4	82.9	17.1

S3Q2	*20*	Stop tail docking (tail-biting resolved)	50.0	50.0	47.6	52.4	55.3	44.7	19.5	80.5
S3Q3	*21*	Management changes needed (tail-biting ongoing)	23.7	76.3	26.8	73.2	28.9	71.1	22.0	78.0
S3Q4	*22*	Compliant (no tail-biting occurring)	9.2	90.8	6.1	93.9	6.6	93.4	6.1	93.9

Supplementary file S2: Raw responses for Control and Training group. For clarity, results for Sections 1, 2 and 4 are shown for Iteration 2 only. For Section 3, results represent binary coding (correct/incorrect) of multiple choice response and both iterations are presented (N.B. the percentages shown are those used in calculating McNemar's test values; they are not directly comparable with Figure 2, which illustrates the percentage of participants who selected each (non-exclusive) answer